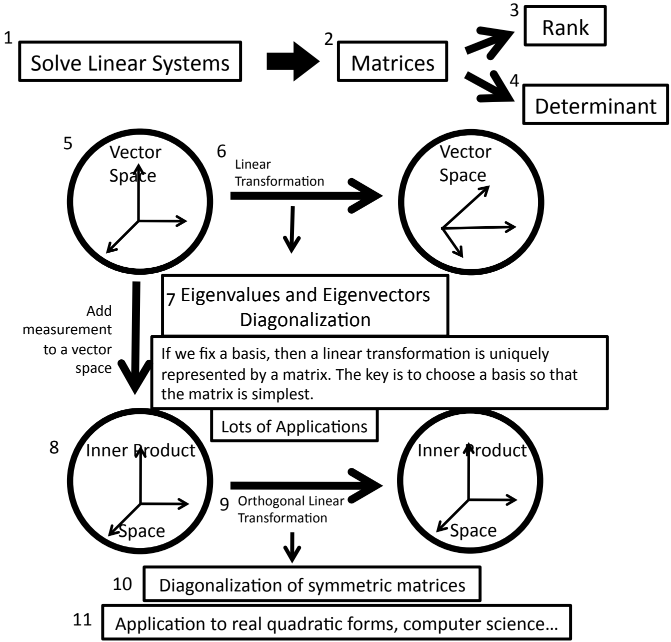
Lecture 1: Linear Algebra Review

Prof. Weiqing Gu

Math 142:
Differential Geometry

A Big Picture of Linear Algebra



Solving Linear Systems

Summary of Key Points

Theorem

Every nonzero $m \times n$ matrix A is row equivalent to a matrix in row echelon form (or to a matrix in reduced row echelon form).

Summary of Key Points

Theorem

Every nonzero $m \times n$ matrix A is row equivalent to a matrix in row echelon form (or to a matrix in reduced row echelon form).

Solving Linear Systems

Consider a linear system of m equations and n unknowns. Let $\overline{A}$ be the augmented matrix representing this system and use elementary operations to change $\overline{A}$ to:

$$\begin{aligned} \text{row echelon form} &\leftrightarrow \text{Gaussian elimination} \\ \text{reduced row echelon form} &\leftrightarrow \text{Gauss-Jordan reduction} \end{aligned}$$

Summary of Key Points

In general,

$$\overline{A} = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right]$$

$$\xrightarrow[\text{operations}]{\text{By elem. row}} \left[\begin{array}{ccccccc|c} 1 & c_{12} & c_{13} & \cdots & & & c_{1n} & d_1 \\ 0 & 0 & 1 & c_{24} & \cdots & & c_{2n} & d_2 \\ \vdots & \vdots & \vdots & \vdots & & & \vdots & \vdots \\ 0 & 0 & \cdots & & 0 & 1 & c_{(k-1)n} & dk-1 \\ 0 & \cdots & & & \vdots & 0 & 1 & d_k \\ 0 & \cdots & & & & \vdots & 0 & d_{k+1} \\ \vdots & & & & & & \vdots & \vdots \\ 0 & \cdots & & & & & 0 & d_m \end{array}\right]$$

Summary of Key Points

This augmented matrix represents the linear system

$$
\begin{array}{ccccccc}
x_1 & + c_{12}x_2 & + c_{13}x_3 & + \cdots & & + c_{1n}x_n & = & d_1 \\
 & & x_3 & + c_{24}x_4 & + \cdots & + c_{2n}x_n & = & d_2 \\
 & & & \vdots & & & & \\
 & & & & x_{n-1} & + c_{(k-1)n}x_n & = & d_{k-1} \\
 & & & & & x_n & = & d_k \\
0x_1 & + & & \cdots & + & 0x_n & = & d_{k+1} \\
\vdots & & & \vdots & & \vdots & & \\
0x_1 & + & & \cdots & + & 0x_n & = & d_m.
\end{array}
$$

Summary of Key Points

By renaming variables we can obtain the linear system

$$
\begin{array}{ccccccccc}
x_1 & + c_{12}x_2 & + \cdots & + c_{1r}x_r & + \cdots & + c_{1n}x_n & = & d_1 \\
 & x_2 & + \cdots & + c_{2r}x_r & + \cdots & + c_{2n}x_n & = & d_2 \\
 & & \ddots & \vdots & & & \vdots & \\
 & & & x_r & + \cdots & + c_{rn}x_n & = & d_r \\
 & & & & & 0 & = & d_{r+1} \\
 & & & & & 0 & = & 0 \\
 & & & & & & \vdots & \\
 & & & & & 0 & = & 0.
\end{array}
$$

Summary of Key Points

By renaming variables we can obtain the linear system

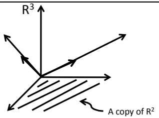
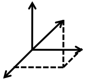
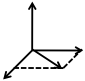
$$\begin{array}{ccccccccc}
x_1 & + c_{12}x_2 & + \cdots & + c_{1r}x_r & + \cdots & + c_{1n}x_n & = & d_1 \\
 & x_2 & + \cdots & + c_{2r}x_r & + \cdots & + c_{2n}x_n & = & d_2 \\
 & & \ddots & \vdots & & & \vdots & \\
 & & & x_r & + \cdots & + c_{rn}x_n & = & d_r \\
 & & & & & 0 & = & d_{r+1} \\
 & & & & & 0 & = & 0 \\
 & & & & & & \vdots & \\
 & & & & & 0 & = & 0.
\end{array}$$

- No solution if $0 \neq d_{r+1}$ (i.e. $\operatorname{rank} \overline{A} \neq \operatorname{rank} A$)
- Unique solution if $0 = d_{r+1}$ and $r = n$.
- Infinitely many solutions if $0 = d_{r+1}$ and $r < n$

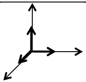
Outline

- "Mimic": A key idea in linear algebra (in fact, in all math)
  1. Understand one important mathematical object very well
  2. Find the characteristic properties of such an object
  3. Use these properties to define a family of objects satisfying these generalized properties
- Examples
  1. Mimic $\mathbb{R}^n \xrightarrow{\text{define}}$ vector spaces
  2. Mimic dot product $\xrightarrow{\text{define}}$ inner products
- Norm, length, angle, distance, unit vector, orthogonality
- Orthogonal projection

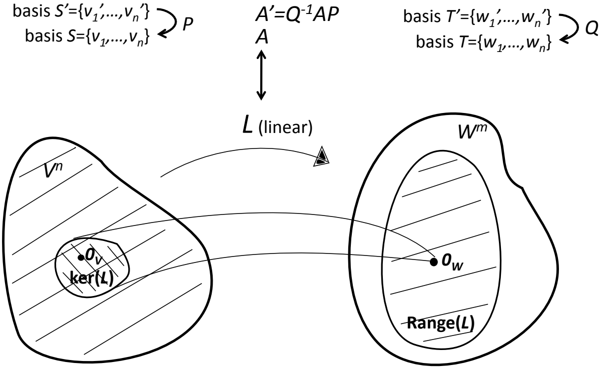
A Big Picture of a Vector Space

Authentic $\mathbb{R}^n$ **Geometrically**	$\xrightarrow{\text{define}}$	Imitation of $\mathbb{R}^n$ **Abstractly**
 		*subspace*
		*linear combination* and *span*
		*linear dependence* and *independence*

A Big Picture of a Vector Space

Authentic $\mathbb{R}^n$ **Geometrically**	$\xrightarrow{\text{define}}$	Imitation of $\mathbb{R}^n$ **Abstractly**
		*basis* and *dimension*
$v = x\vec{i} + y\vec{j} + z\vec{k} \leftrightarrow (x, y, z)$		*coordinates*
$\mathbb{R}^n$	$\xleftrightarrow{\text{isomorphic}}$	n-dimensional vector space

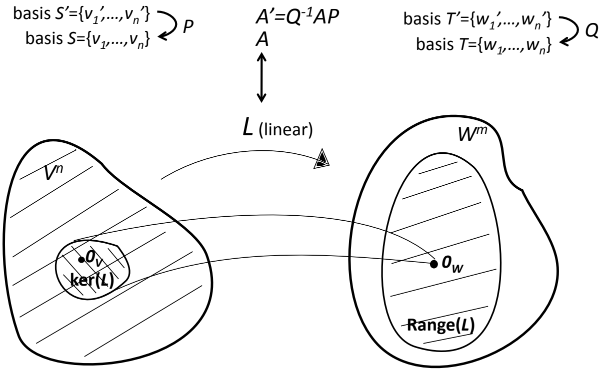
Linear Transformations and Matrices



ker(L) is a subspace of V^n

Range(L) is a subspace of W^m

Linear Transformations and Matrices



ker(L) is a subspace of V^n

Range(L) is a subspace of W^m

$$\begin{array}{ccccc} \dim \ker L & + & \dim \operatorname{range} L & = & \dim V^n \\ \updownarrow & & \updownarrow & & \updownarrow \\ \text{nullity } A & + & \text{rank } A & = & n \end{array}$$

Linear Transformations and Matrices

Encodings

Let $\mathcal{L}(V^n, V^m)$ be the set of all linear transformations $L : V^n \rightarrow V^m$ and $\mathcal{M}(m, n)$ be the set of all $m \times n$ matrices. Then

$$\begin{aligned} \mathcal{L}(V^m, V^n) &\cong \mathcal{M}(m, n) \\ L_1 + L_2 &\leftrightarrow A_1 + A_2 \\ cL &\leftrightarrow cA \\ L_1 L_2 &\leftrightarrow A_1 A_2 \end{aligned}$$

Kernel of a Linear Transformation

Definition

Let $L : V \to W$ be a linear transformation of a vector space V into a vector space W. The **kernel** of L, denoted $\ker L$, is the subset of V consisting of all elements $\mathbf{v}$ of V such that $L\mathbf{v} = \mathbf{0}_W$.

Kernel of a Linear Transformation

Definition

Let $L : V \to W$ be a linear transformation of a vector space V into a vector space W. The **kernel** of L, denoted $\ker L$, is the subset of V consisting of all elements $\mathbf{v}$ of V such that $L\mathbf{v} = \mathbf{0}_W$.

Definition

A linear transformation $L : V \to W$ is called **one-to-one** if it is a one to one function; that is, if $\mathbf{v}_1 \neq \mathbf{v}_2$ implies that $L\mathbf{v}_1 \neq L\mathbf{v}_2$. An equivalent statement is that L is one-to-one if $L\mathbf{v}_1 = L\mathbf{v}_2$ implies that $\mathbf{v}_1 = \mathbf{v}_2$.

Kernel of a Linear Transformation

Definition

Let $L : V \to W$ be a linear transformation of a vector space V into a vector space W. The **kernel** of L, denoted $\ker L$, is the subset of V consisting of all elements $\mathbf{v}$ of V such that $L\mathbf{v} = \mathbf{0}_W$.

Definition

A linear transformation $L : V \to W$ is called **one-to-one** if it is a one to one function; that is, if $\mathbf{v}_1 \neq \mathbf{v}_2$ implies that $L\mathbf{v}_1 \neq L\mathbf{v}_2$. An equivalent statement is that L is one-to-one if $L\mathbf{v}_1 = L\mathbf{v}_2$ implies that $\mathbf{v}_1 = \mathbf{v}_2$.

Theorem

Let $L : V \to W$ be a linear transformation of a vector space V into a vector space W. Then

(a) $\ker L$ *is a subspace of V.*

(b) *L is one-to-one if and only if $\ker L = \{\mathbf{0}_V\}$ if and only if* $\dim \ker L = 0$.

Eigenvalues and Eigenvectors

Goal

Let $L : V \to V$ be a linear transformation of an n-dimensional vector space V into itself. We want to find a basis T of V such that the matrix representation D with respect to T is in the simplest possible form.

Eigenvalues and Eigenvectors

Goal

Let $L : V \to V$ be a linear transformation of an n-dimensional vector space V into itself. We want to find a basis T of V such that the matrix representation D with respect to T is in the simplest possible form.

Hope

To find a basis T of V such that D is a diagonal matrix. This is not always possible—if it is, we say that L is *diagonalizable*.

Eigenvalues and Eigenvectors

Goal

Let $L : V \to V$ be a linear transformation of an n-dimensional vector space V into itself. We want to find a basis T of V such that the matrix representation D with respect to T is in the simplest possible form.

Hope

To find a basis T of V such that D is a diagonal matrix. This is not always possible—if it is, we say that L is *diagonalizable*.

When can we do so?

Eigenvalues and Eigenvectors

Goal

Let $L : V \to V$ be a linear transformation of an n-dimensional vector space V into itself. We want to find a basis T of V such that the matrix representation D with respect to T is in the simplest possible form.

Hope

To find a basis T of V such that D is a diagonal matrix. This is not always possible—if it is, we say that L is *diagonalizable*.

When can we do so?

Why?

For example, let $L : \mathbb{R}^3 \to \mathbb{R}^3$ be given by $L(a_1, a_2, a_3) = (2a_1 - a_3, a_1 + a_2 - a_3, a_3)$, and find L^{100}.

Diagonalization

When can we diagonalize A over $\mathbb{R}$?

When do there exist n linearly independent eigenvectors of A in $\mathbb{R}$?

Diagonalization

When can we diagonalize A over $\mathbb{R}$?

When do there exist n linearly independent eigenvectors of A in $\mathbb{R}$?

1. Find all eigenvalues of A

Diagonalization

When can we diagonalize A over $\mathbb{R}$?

When do there exist n linearly independent eigenvectors of A in $\mathbb{R}$?

1. Find all eigenvalues of A
2. If there is an eigenvalue of A that is not in $\mathbb{R}$, <u>stop</u>! A is not diagonalizable <u>over $\mathbb{R}$</u>.

Diagonalization

When can we diagonalize A over $\mathbb{R}$?

When do there exist n linearly independent eigenvectors of A in $\mathbb{R}$?

1. Find all eigenvalues of A
2. If there is an eigenvalue of A that is not in $\mathbb{R}$, stop! A is not diagonalizable over $\mathbb{R}$.
3. If all the eigenvalues of A (n many, counting multiplicity) are in $\mathbb{R}$, then for each eigenvalue λ, find $\text{rank}(A - \lambda I)$.

Diagonalization

When can we diagonalize A over $\mathbb{R}$?

When do there exist n linearly independent eigenvectors of A in $\mathbb{R}$?

1. Find all eigenvalues of A
2. If there is an eigenvalue of A that is not in $\mathbb{R}$, <u>stop</u>! A is not diagonalizable <u>over $\mathbb{R}$</u>.
3. If all the eigenvalues of A (n many, counting multiplicity) are in $\mathbb{R}$, then for each eigenvalue λ, find $\operatorname{rank}(A - \lambda I)$. If there exists a λ such that

$$\text{multiplicity of } \lambda \neq n - \operatorname{rank}(A - \lambda I) = \operatorname{nullity}(A - \lambda I),$$

<u>stop</u>! A is not diagonalizable over $\mathbb{R}$.

Diagonalization

When can we diagonalize A over $\mathbb{R}$?

When do there exist n linearly independent eigenvectors of A in $\mathbb{R}$?

1. Find all eigenvalues of A
2. If there is an eigenvalue of A that is not in $\mathbb{R}$, stop! A is not diagonalizable over $\mathbb{R}$.
3. If all the eigenvalues of A (n many, counting multiplicity) are in $\mathbb{R}$, then for each eigenvalue λ, find $\operatorname{rank}(A - \lambda I)$. If there exists a λ such that

$$\text{multiplicity of } \lambda \neq n - \operatorname{rank}(A - \lambda I) = \operatorname{nullity}(A - \lambda I),$$

stop! A is not diagonalizable over $\mathbb{R}$.
4. Otherwise, A is diagonalizable over $\mathbb{R}$.

Diagonalization

Why?

Let

$$\begin{aligned} V_\lambda &= \{x \in \mathbb{R}^n \mid Ax = \lambda x\} \\ &= \{x \in \mathbb{R}^n \mid (A - \lambda I)x = 0\} \\ &= \operatorname{null}(A - \lambda I). \end{aligned}$$

Thus,

$$\begin{aligned} \dim V_\lambda &= \dim \operatorname{null}(A - \lambda I) \\ &= n - \operatorname{rank}(A - \lambda I) \\ &= \operatorname{nullity}(A - \lambda I). \end{aligned}$$

Eigenvalues and Eigenvectors

Facts

- $\dim V_\lambda \geq 1$
- $\dim V_\lambda \leq$ the multiplicity of λ (i.e. geometric multiplicity $\leq$ algebraic multiplicity)
- Eigenvectors belonging to different eigenvalues must be linearly independent

Eigenvalues and Eigenvectors

Facts

- $\dim V_\lambda \geq 1$
- $\dim V_\lambda \leq$ the multiplicity of λ (i.e. geometric multiplicity $\leq$ algebraic multiplicity)
- Eigenvectors belonging to different eigenvalues must be linearly independent

Theorem

An $n \times n$ matrix A is similar to a diagonal matrix D if and only if $\mathbb{R}^n$ has a basis of eigenvectors of A. Moreover, the elements on the main diagonal of D are the eigenvalues of A.

Eigenvalues and Eigenvectors

Facts

- $\dim V_\lambda \geq 1$
- $\dim V_\lambda \leq$ the multiplicity of λ (i.e. geometric multiplicity $\leq$ algebraic multiplicity)
- Eigenvectors belonging to different eigenvalues must be linearly independent

Theorem

An $n \times n$ matrix A is similar to a diagonal matrix D if and only if $\mathbb{R}^n$ has a basis of eigenvectors of A. Moreover, the elements on the main diagonal of D are the eigenvalues of A.

Thus, to find n linearly independent eigenvectors, for each λ we must have $\dim V_\lambda = \text{mult}(\lambda)$.

Eigenvalues and Eigenvectors

Facts

- $\dim V_\lambda \geq 1$
- $\dim V_\lambda \leq$ the multiplicity of λ (i.e. geometric multiplicity $\leq$ algebraic multiplicity)
- Eigenvectors belonging to different eigenvalues must be linearly independent

Theorem

An $n \times n$ matrix A is similar to a diagonal matrix D if and only if $\mathbb{R}^n$ has a basis of eigenvectors of A. Moreover, the elements on the main diagonal of D are the eigenvalues of A.

Thus, to find n linearly independent eigenvectors, for each λ we must have $\dim V_\lambda = \text{mult}(\lambda)$.

Theorem

An $n \times n$ matrix A is diagonalizable if all the roots of its characteristic polynomial are real and distinct.

Eigenvalues and Eigenvectors

In the case that A is diagonalizable, to find P such that $P^{-1}AP = D$ we need

1. For each λ_i, find a basis for V_{λ_i}, i.e. solve

$$(A - \lambda_i I)x = 0$$

   and find a basis for $\operatorname{null}(A - \lambda_i I)$. Say you find a basis $x_{i1}, \ldots, x_{ik_i}$, where k_i is the multiplicity of λ_i.

2. Let $P = (x_{11}, \ldots, x_{1k_1}, \ldots, x_{i1}, \ldots, x_{ik_i}, \ldots, x_{r1}, \ldots, x_{rk_r})$. Then

$$P^{-1}AP = \begin{bmatrix} \lambda_1 I_{k_1} & & & & \\ & \ddots & & & \\ & & \lambda_i I_{k_i} & & \\ & & & \ddots & \\ & & & & \lambda_r I_{k_r} \end{bmatrix}.$$

Inner Product Spaces

Big Ideas

- Standard inner product space $\mathbb{R}^2$, $\mathbb{R}^3$
- Inner product space
- Matrix of an inner product
- Orthonormal bases
- Gram-Schmidt process

The Dot Product

Key

The standard inner product (or dot product) on $\mathbb{R}^3$ (or $\mathbb{R}^n$). If we define $u \cdot v = u_1 v_1 + u_2 v_2 + u_3 v_3$, then every measurement can be expressed in terms of the dot product.

The Dot Product

Key

The standard inner product (or dot product) on $\mathbb{R}^3$ (or $\mathbb{R}^n$). If we define $u \cdot v = u_1 v_1 + u_2 v_2 + u_3 v_3$, then every measurement can be expressed in terms of the dot product.

Let $u = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$ and $v = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$.

- Length of u: $\|u\| = \sqrt{u_1^2 + u_2^2 + u_3^2} = \sqrt{u \cdot u}$
- Distance between u and v:
  $d(u, v) = \|u - v\| = \sqrt{(u - v) \cdot (u - v)}$
- Angle between u and v:

$$\cos\theta = \frac{u \cdot v}{\|u\|\|v\|}, \qquad 0 \leq \theta < 2\pi.$$

- Orthogonality: $u \cdot v = 0$.

Mimicking the Dot Product

For an arbitrary vector space V, we also want to define measurements.

Key

We need to define the "inner product" for V. We want to imitate the characteristic properties in $\mathbb{R}^3$.

Mimicking the Dot Product

For an arbitrary vector space V, we also want to define measurements.

Key

We need to define the "inner product" for V. We want to imitate the characteristic properties in $\mathbb{R}^3$.

Characteristic Properties of the Dot Product

Let $u, v, w \in \mathbb{R}^3$, $c \in \mathbb{R}$.

(a) $u \cdot u > 0$ if $u \neq 0$; $u \cdot u = 0$ if and only if $u = 0$

(b) $u \cdot v = v \cdot u$

(c) $(u + v) \cdot w = u \cdot w + v \cdot w$

(d) $(cu) \cdot v = c(u \cdot v)$

Mimicking the Dot Product

For an arbitrary vector space V, we also want to define measurements.

Key

We need to define the "inner product" for V. We want to imitate the characteristic properties in $\mathbb{R}^3$.

Definition

Let V be any real vector space. An **inner product** on V is a function that assigns to each ordered pair of vectors u, v in V a real number $\langle u, v\rangle$ satisfying

(a) $\langle u, u\rangle > 0$ for $u \neq 0$; $\langle u, u\rangle = 0$ if and only if $u = 0$

(b) $\langle v, u\rangle = \langle u, v\rangle$ for any $u, v \in V$

(c) $\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$ for any $u, v, w \in V$

(d) $\langle cu, v\rangle = c\langle u, v\rangle$ for $u, v \in V$ and $c \in \mathbb{R}$.

Inner Product Measurements

- Length of u: $\|u\| = \sqrt{\langle u, u\rangle}$
- Distance between u and v: $d(u, v) = \|u - v\|$
- Angle between u and v: $\cos\theta = \dfrac{\langle u, v\rangle}{\|u\|\|v\|}$ (Note: must show $|\langle u, v\rangle| \leq \|u\|\|v\|$ before defining $\cos\theta$)
- Orthogonality: $\langle u, v\rangle = 0$

Inner Product Measurements

- Length of u: $\|u\| = \sqrt{\langle u, u \rangle}$
- Distance between u and v: $d(u, v) = \|u - v\|$
- Angle between u and v: $\cos\theta = \dfrac{\langle u, v \rangle}{\|u\|\|v\|}$ (Note: must show $|\langle u, v \rangle| \leq \|u\|\|v\|$ before defining $\cos\theta$)
- Orthogonality: $\langle u, v \rangle = 0$

Theorem (Cauchy-Schwartz Inequality)

If u and v are any two vectors in an inner product space V, then

$$\langle u, v \rangle^2 \leq \|u\|^2 \|v\|^2.$$

Theorem (Triangle Inequality)

If u and v are any vectors in an inner product space V, then $\|u + v\| \leq \|u\| + \|v\|$.

Matrix of an Inner Product

Theorem

Let $S = \{u_1, u_2, \ldots u_n\}$ be an ordered basis for a finite-dimensional vector space V and assume that we are given an inner product on V. Let $c_{ij} = \langle u_i, u_j \rangle$, and $C = [c_{ij}]$. Then

(a) *C is a symmetric, positive-definite matrix.*

(b) *C determines $\langle v, w \rangle$ for every v and w in V.*

Orthonormal Bases and Orthogonal Matrices

Definition

A real vector space that has an inner product defined on it is called an *inner product space*. If the space is finite-dimensional, it is called a *Euclidean space*.

Orthonormal Bases and Orthogonal Matrices

Definition

A real vector space that has an inner product defined on it is called an *inner product space*. If the space is finite-dimensional, it is called a *Euclidean space*.

Definition

Let V be an inner product space. A set S of vectors in V is called orthogonal if any two distinct vectors in S are orthogonal. If, in addition, each vector in S is of unit length, then S is called orthonormal.

Orthonormal Bases and Orthogonal Matrices

Definition

A real vector space that has an inner product defined on it is called an *inner product space*. If the space is finite-dimensional, it is called a *Euclidean space*.

Definition

Let V be an inner product space. A set S of vectors in V is called orthogonal if any two distinct vectors in S are orthogonal. If, in addition, each vector in S is of unit length, then S is called orthonormal.

Theorem

Let $S = \{u_1, u_2, \ldots, u_n\}$ be a finite orthogonal set of nonzero vectors in an inner product space. Then S is linearly independent.

Orthonormal Bases and Orthogonal Matrices

In a Euclidean space V, we are interested in an

Orthonormal Basis

Let $S = \{v_1, v_2, \ldots, v_n\} \subset V$. We say that S is an *orthonormal basis* if (i) S is a basis for V, and (ii) S is an orthonormal set. Let $A = \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix}$. Then $A^T A = I$.

Orthogonal Matrix

A square matrix is orthogonal if $A^T A = I$.

Outline

- An example of how to use the Gram-Schmidt method to turn an arbitrary basis into an orthonormal one
- An example of how to obtain a QR-factorization of a nonsingular matrix
- Extend an orthonormal set to an orthonormal basis of $\mathbb{R}^n$
- Summaries on Gram-Schmidt QR factorization

The Gram-Schmidt Process

Theorem

For every Euclidean space V, we can obtain an orthonormal basis S.

The Gram-Schmidt Process

Theorem

For every Euclidean space V, we can obtain an orthonormal basis S.

First, we need some ideas on projection. Then we'll see why Gram-Schmidt makes sense.

The Gram-Schmidt Process

Theorem

For every Euclidean space V, we can obtain an orthonormal basis S.

First, we need some ideas on projection. Then we'll see why Gram-Schmidt makes sense.

Example

Let $V = \mathbb{R}^3$. Let $S = \left\{ v_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, v_2 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, v_3 = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} \right\}$ be a basis of $\mathbb{R}^3$. Find an orthonormal basis $\{u_1, u_2, u_3\}$ that satisfies

(i) u_1 is parallel to v_1, and

(ii) u_2 lies in the plane of v_1 and v_2.

Process

Step 1: Get u_1

$$u_1 = \frac{v_1}{\|v_1\|}$$

Step 2: Get u_2

$$w_2 = v_2 - \langle v_2, u_1 \rangle u_1, \qquad u_2 = \frac{w_2}{\|w_2\|}$$

Step 3: Get u_3

$$w_3 = v_3 - \langle v_3, u_1 \rangle u_1 - \langle v_3, u_2 \rangle u_2, \qquad u_3 = \frac{w_3}{\|w_3\|}$$

Process

Step 1: Get u_1

$$u_1 = \frac{v_1}{\|v_1\|}$$

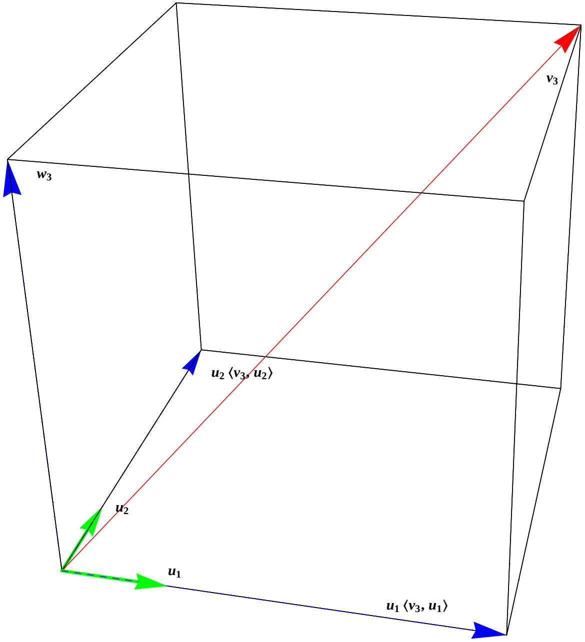
Step 2: Get u_2

$$w_2 = v_2 - \langle v_2, u_1 \rangle u_1, \qquad u_2 = \frac{w_2}{\|w_2\|}$$

Step 3: Get u_3

$$w_3 = v_3 - \langle v_3, u_1 \rangle u_1 - \langle v_3, u_2 \rangle u_2, \qquad u_3 = \frac{w_3}{\|w_3\|}$$

This process is called the Gram-Schmidt process and can be used to convert an independent set $\{v_1, v_2, \ldots, v_n\}$ into an orthonormal set $\{u_1, u_2, \ldots, u_n\}$ so that each u_i depends only on $\{v_1, v_2, \ldots, v_i\}$.



v_3
w_3
$u_2 \langle v_3, u_2 \rangle$
u_2
u_1
$u_1 \langle v_3, u_1 \rangle$

QR Factorization of a Matrix

Following the above process, we can write

$$
\begin{aligned}
v_1 &= \|v_{11}\|u_1 & &= r_1 1 u_1 \\
v_2 &= \langle v_2, u_1\rangle u_1 \quad + \|w_2\|w_2 & &= r_{12}u_1 + r_{22}u_2 \\
v_3 &= \langle v_3, u_1\rangle u_1 \quad + \langle v_3, u_2\rangle u_2 + \|w_3\|w_3 & &= r_{13}u_1 + r_{23}u_2 + r_{33}u_3,
\end{aligned}
$$

where we define $r_{ij} = \langle v_j, u_i\rangle$ if $i \neq j$ and $r_{ii} = \|w_i\| \neq 0$.

QR Factorization of a Matrix

Following the above process, we can write

$$
\begin{aligned}
v_1 &= \|v_{11}\| u_1 & &= r_1 1 u_1 \\
v_2 &= \langle v_2, u_1 \rangle u_1 & + \|w_2\| w_2 & = r_{12} u_1 + r_{22} u_2 \\
v_3 &= \langle v_3, u_1 \rangle u_1 & + \langle v_3, u_2 \rangle u_2 + \|w_3\| w_3 & = r_{13} u_1 + r_{23} u_2 + r_{33} u_3,
\end{aligned}
$$

where we define $r_{ij} = \langle v_j, u_i \rangle$ if $i \neq j$ and $r_{ii} = \|w_i\| \neq 0$. Thus,

$$
\begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix} = \underbrace{\begin{bmatrix} u_1 & u_2 & u_3 \end{bmatrix}}_{Q} \underbrace{\begin{bmatrix} r_{11} & r_{12} & r_{13} \\ 0 & r_{22} & r_{23} \\ 0 & 0 & r_{33} \end{bmatrix}}_{R}
$$

QR Factorization of a Matrix

Following the above process, we can write

$$\begin{aligned}
v_1 &= \|v_{11}\| u_1 & &= r_1 1 u_1 \\
v_2 &= \langle v_2, u_1\rangle u_1 + \|w_2\| w_2 & &= r_{12} u_1 + r_{22} u_2 \\
v_3 &= \langle v_3, u_1\rangle u_1 + \langle v_3, u_2\rangle u_2 + \|w_3\| w_3 & &= r_{13} u_1 + r_{23} u_2 + r_{33} u_3,
\end{aligned}$$

where we define $r_{ij} = \langle v_j, u_i\rangle$ if $i \neq j$ and $r_{ii} = \|w_i\| \neq 0$. Thus,

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} = \underbrace{\begin{bmatrix} \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{2}{\sqrt{6}} & 0 \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \end{bmatrix}}_{Q} \underbrace{\begin{bmatrix} \sqrt{3} & \frac{4}{\sqrt{3}} & \frac{4}{\sqrt{3}} \\ 0 & \frac{\sqrt{6}}{3} & -\frac{1}{\sqrt{6}} \\ 0 & 0 & \frac{1}{\sqrt{2}} \end{bmatrix}}_{R}$$

Diagonalization of Symmetric Matrices

Every symmetric matrix can be diagonalized over $\mathbb{R}$!

Diagonalization of Symmetric Matrices

Every symmetric matrix can be diagonalized over $\mathbb{R}$!

Moreover, there exists an orthogonal matrix P such that $P^{-1}AP = P^TAP = D$, a diagonal matrix.

Diagonalization of Symmetric Matrices

Every symmetric matrix can be diagonalized over $\mathbb{R}$!

Moreover, there exists an orthogonal matrix P such that $P^{-1}AP = P^TAP = D$, a diagonal matrix.

WHY?

Diagonalization of Symmetric Matrices

Key 1

Theorem 6.6: *All the roots of the characteristic polynomial of a real symmetric matrix are real numbers.*

Key 2

If the symmetric matrix A has an eigenvalue λ of multiplicity k, then the solution space of the homogeneous system $(\lambda I_n - A)\mathbf{x} = 0$ has dimension k. This means that there exist k linearly independent eigenvectors of A associated with the eigenvalue λ.

Diagonalization of Symmetric Matrices

Key 3

Theorem 6.7: *If A is a symmetric $n \times n$ matrix, then eigenvectors that belong to distinct eigenvalues of A are orthogonal.*

Key 4

By the Gram-Schmidt process we can choose an orthonormal basis for this solution space. Thus we obtain a set of k orthonormal eigenvectors associated with the eigenvalue λ. Since eigenvectors associated with distinct eigenvalues are orthogonal, if we form the set of all eigenvectors we get an orthonormal set. Hence, the matrix P whose columns are the eigenvectors is orthogonal.

Applications of Eigenvalues to Geometry

Real Quadratic Forms

In precalculus and calculus courses you have seen that the graph of the equation

$$ax^2 + 2bxy + cy^2 = d,$$

where a, b, c, and d are real numbers, is a *conic section* centered at the origin of a rectangular Cartesian coordinate system in two-dimensional space.

Applications of Eigenvalues to Geometry

Real Quadratic Forms

In precalculus and calculus courses you have seen that the graph of the equation

$$ax^2 + 2bxy + cy^2 = d,$$

where a, b, c, and d are real numbers, is a *conic section* centered at the origin of a rectangular Cartesian coordinate system in two-dimensional space. Similarly, the graph of the equation

$$ax^2 + 2dxy + 2exz + by^2 + 2fyz + cz^2 = g,$$

where a, b, c, d, e, f, and g are real numbers, is a *quadratic surface* centered at the origin of a rectangular Cartesian coordinate system in three-dimensional space.

Applications of Eigenvalues to Geometry

Real Quadratic Forms

In precalculus and calculus courses you have seen that the graph of the equation

$$ax^2 + 2bxy + cy^2 = d,$$

where a, b, c, and d are real numbers, is a *conic section* centered at the origin of a rectangular Cartesian coordinate system in two-dimensional space. Similarly, the graph of the equation

$$ax^2 + 2dxy + 2exz + by^2 + 2fyz + cz^2 = g,$$

where a, b, c, d, e, f, and g are real numbers, is a *quadratic surface* centered at the origin of a rectangular Cartesian coordinate system in three-dimensional space.

We can write these equations into matrix form.

Real Quadratic Forms

Example

Write the following quadratic form into matrix form:

$$3x^2 - 7xy + 5xz + 4y^2 - 4yz - 3z^2.$$

Real Quadratic Forms

Example

Write the following quadratic form into matrix form:

$$3x^2 - 7xy + 5xz + 4y^2 - 4yz - 3z^2.$$

Example

The following expressions are quadratic forms:

(a) $3x^2 - 5xy - 7y^2 = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 3 & -\frac{5}{2} \\ -\frac{5}{2} & -7 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$

(b) $3x^2 - 7xy + 5xz + 4y^2 - 4yz - 3z^2 =$
$$\begin{bmatrix} x & y & z \end{bmatrix} \begin{bmatrix} 3 & -\frac{7}{2} & \frac{5}{2} \\ -\frac{7}{2} & 4 & -2 \\ \frac{5}{2} & -2 & -3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

Real Quadratic Forms

Example

Identify the graph of the equation

$$2x^2 + 6xy + 2y^2 = 5.$$

Real Quadratic Forms

Example

Identify the graph of the equation

$$2x^2 + 6xy + 2y^2 = 5.$$

Example

Classify the quadratic surface

$$4xy + 4yz + 4xz = 7.$$

Real Quadratic Forms

Example

Identify the surface given by $4xy + 4xz + 4yz = 7$.

Real Quadratic Forms

Example

Identify the surface given by $4xy + 4xz + 4yz = 7$.

A Linear Algebra Solution

Let $A = \begin{bmatrix} 0 & 2 & 2 \\ 2 & 0 & 2 \\ 2 & 2 & 0 \end{bmatrix}$.

Real Quadratic Forms

Example

Identify the surface given by $4xy + 4xz + 4yz = 7$.

A Linear Algebra Solution

Let $A = \begin{bmatrix} 0 & 2 & 2 \\ 2 & 0 & 2 \\ 2 & 2 & 0 \end{bmatrix}$. Since A is symmetric, we can find an orthogonal matrix P such that $P^T AP$ is diagonal.

Real Quadratic Forms

Example

Identify the surface given by $4xy + 4xz + 4yz = 7$.

A Linear Algebra Solution

Let $A = \begin{bmatrix} 0 & 2 & 2 \\ 2 & 0 & 2 \\ 2 & 2 & 0 \end{bmatrix}$. Since A is symmetric, we can find an orthogonal matrix P such that $P^T AP$ is diagonal. A has eigenvalues $\lambda_1 = -2$ (multiplicity 2) and $\lambda_2 = 4$. The associated eigenvectors are

$$\underbrace{v_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \quad v_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}}_{\lambda_1}, \quad \underbrace{v_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}}_{\lambda_2}.$$

Real Quadratic Forms

A Linear Algebra Solution Continued

Applying the Gram-Schmidt process, we obtain the orthonormal basis

$$u_1 = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}, \quad u_2 = \begin{bmatrix} -\frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \end{bmatrix}, \quad u_3 = \begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}.$$

Thus, we know

$$P = \begin{bmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{bmatrix},$$

and $P^{-1}AP = P^TAP = \begin{bmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 4 \end{bmatrix}$.

Real Quadratic Forms

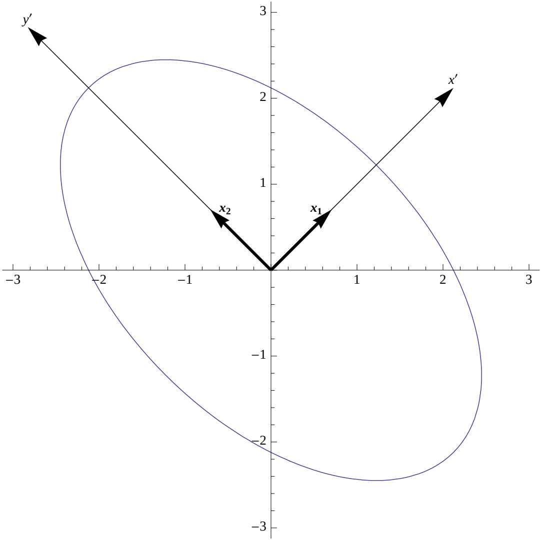
Example

Consider the conic section whose equation is

$$g(\mathbf{x}) = 2x^2 + 2xy + 2y^2 = 9.$$

What kind of curve is this? Geometric meaning:

Real Quadratic Forms



Real Quadratic Forms

Example

Consider the conic section whose equation is

$$g(\mathbf{x}) = 2x^2 + 2xy + 2y^2 = 9.$$

What kind of curve is this? Geometric meaning:

In engineering, one often deals with symmetric matrices (e.g. mass matrix, stiffness matrix). For this reason, finding eigenvectors is a matter of finding rotation axes or saying that it is a simple rotation.

Real Quadratic Forms

Definition

If A is a symmetric matrix, then the function $g : \mathbb{R}^n \to \mathbb{R}$ (a real valued function on $\mathbb{R}^n$) defined by

$$g(\mathbf{x}) = \mathbf{x}^T A\mathbf{x},$$

where

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix},$$

is called a *real quadratic form in the n variables $x_1, x_2, \ldots, x_n$*. The matrix A is called the *matrix of the quadratic form g*. We shall also denote the quadratic form by $g(\mathbf{x})$.

Real Quadratic Forms

Definition

If A and B are $n \times n$ matrices, we say that B is *congruent* to A if $B = P^T AP$ for a nonsingular matrix P.

Real Quadratic Forms

Definition

If A and B are $n \times n$ matrices, we say that B is *congruent* to A if $B = P^T AP$ for a nonsingular matrix P.

Definition

Two quadratic forms g and h with matrices A and B, respectively are said to be *equivalent* if A and B are congruent.

Real Quadratic Forms

Definition

If A and B are $n \times n$ matrices, we say that B is *congruent* to A if $B = P^T AP$ for a nonsingular matrix P.

Definition

Two quadratic forms g and h with matrices A and B, respectively are said to be *equivalent* if A and B are congruent.

Theorem (Principal Axes Theorem)

Any quadratic form in n variables $g(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ is equivalent by means of an orthogonal matrix P to a quadratic form $h(\mathbf{y}) = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2$, where

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}.$$

Real Quadratic Forms

Theorem

A quadratic form $g(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ is equivalent to a quadratic form

$$h(\mathbf{y}) = y_1^2 + y_2^2 + \cdots + y_p^2 - y_{p+1}^2 - y_{p+2}^2 - \cdots - y_r^2.$$

Theorem

A symmetric matrix A is positive definite if and only if all the eigenvalues of A are positive.

A quadratic form is then called *positive definite* if its matrix is positive definite.

Real Quadratic Forms

Example

Consider the quadratic form g in the variables x, y, and z, defined by

$$g(\mathbf{x}) = 2x^2 + 4y^2 + 6yz - 4z^2.$$

Determine a quadratic form h of the forms in the theorems above to which g is equivalent. Is g positive definite?

Real Quadratic Forms

Example

Consider the quadratic form g in the variables x, y, and z, defined by

$$g(\mathbf{x}) = 2x^2 + 4y^2 + 6yz - 4z^2.$$

Determine a quadratic form h of the forms in the theorems above to which g is equivalent. Is g positive definite?

Solution

The matrix of g is

$$A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 4 & 3 \\ 0 & 3 & -4 \end{bmatrix},$$

and the eigenvalues of A are $\lambda_1 = 2$, $\lambda_2 = 5$, and $\lambda_3 = -5$. Let h be the quadratic form in the variables x', y', and z' defined by $h(\mathbf{y}) = 2x'^2 + 5y'^2 - 5z'^2$. Then g and h are equivalent by means of some orthogonal matrix.

Real Quadratic Forms

Definition

Let $g(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ be a quadratic form in n variables. The difference between the number of positive eigenvalues of A and the number of negative eigenvalues of A is called the *signature* of the quadratic form g.

Real Quadratic Forms

Definition

Let $g(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ be a quadratic form in n variables. The difference between the number of positive eigenvalues of A and the number of negative eigenvalues of A is called the *signature* of the quadratic form g.

Fact

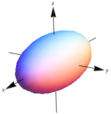
Two quadratic forms g and h are equivalent if and only if they have equal ranks and signatures.

Example

$$g(\mathbf{x}) = x_1^2 + x_2^2 + x_3^2 + 2x_1x_2$$
$$h(\mathbf{x}) = 2x_2^2 + 2x_3^2 + 2x_2x_3$$

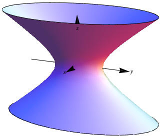
Ellipsoid

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$$



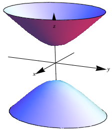
Hyperboloid of one sheet

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1$$



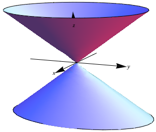
Hyperboloid of two sheets

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = -1$$



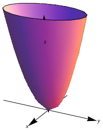
Elliptic cone

$$z^2 = \frac{x^2}{a^2} + \frac{y^2}{b^2}$$



Elliptic Paraboloid

$$z = \frac{x^2}{a^2} + \frac{y^2}{b^2}$$



Elliptic Paraboloid

$$z = \frac{x^2}{a^2} - \frac{y^2}{b^2}$$